

1 DAVID H. KRAMER, SBN 168452
Email: dkramer@wsgr.com
2 MAURA L. REES, SBN 191698
Email: mrees@wsgr.com
3 WILSON SONSINI GOODRICH & ROSATI
Professional Corporation
4 650 Page Mill Road
Palo Alto, CA 94304-1050
5 Telephone: (650) 493-9300

6 ERIC P. TUTTLE, SBN 248440
Email: eric.tuttle@wsgr.com
7 WILSON SONSINI GOODRICH & ROSATI
Professional Corporation
8 701 Fifth Avenue, Suite 5100
Seattle, WA 98104-7036
9 Telephone: (206) 883-2500

10 PAUL J. SAMPSON, *pro hac vice*
Email: psampson@wsgr.com
11 WILSON SONSINI GOODRICH & ROSATI
Professional Corporation
12 95 S State Street, Suite 1000
Salt Lake City, UT 84111
13 Telephone: (801) 401-8510

14 *Counsel for Defendants*
GOOGLE LLC AND ALPHABET INC.

15
16 **UNITED STATES DISTRICT COURT**
17 **NORTHERN DISTRICT OF CALIFORNIA**
18 **SAN JOSE DIVISION**
19

20 *In re Google Generative AI Copyright*
21 *Litigation*

Master File Case No.: 5:23-cv-03440-EKL
Consolidated with Case No.: 5:24-cv-02531-EKL

22 **DEFENDANTS' RESPONSE TO**
23 **PLAINTIFFS' STATEMENT RE**
DISCOVERY DISPUTE RAISED
24 **AT ECF No. 120-2**

25 Referral: Hon. Susan van Keulen, U.S.M.J.
26
27
28

1 Pursuant to Judge Van Keulen’s April 8, 2025 order, ECF No. 121, Defendants’ respond to
2 the discovery dispute raised by Plaintiffs at ECF No. 120-2 as follows:

3 Shortly after serving their first discovery requests, Plaintiffs demanded a Rule 30(b)(6)
4 deposition of Defendant Google that covers essentially every issue in the case. If the goal were a
5 truly productive deposition, Plaintiffs would seek one after Google produced its documents and
6 would afford Google sufficient time to identify and prepare appropriate witnesses on specific
7 topics. Instead, on the pretense that Plaintiffs can meaningfully cover 10 sweeping topics in 90
8 minutes, Plaintiffs insist the deposition must take place immediately, without sufficient time for
9 witness preparation or document review. Google cannot adequately prepare, and Plaintiffs cannot
10 legitimately question, a witness on “Google’s doctrinal positions on fair use,” one of the
11 “subtopics” at issue in this motion. *See* Ex. 1 at 2. Likewise, Google cannot be expected to present
12 corporate testimony on the vague and overbroad topic of its “Data Usage Policies And Licensing.”
13 *Id.*

14 This deposition will not help Plaintiffs find facts; it will only help them to play a game of
15 gotcha, hoping to catch a witness unprepared, take testimony out of context, and foment future
16 discovery disputes. Moreover, like many of Plaintiffs’ early discovery demands, their pursuit of
17 this deposition disregards Judge Lee’s direction that they “really focus on discovery in terms of
18 what is needed for the class cert to begin with.” ECF No. 93 at 21:19-21. Plaintiffs’ motion to
19 compel this burdensome, wasteful and premature deposition should be denied.

20 **Procedural Background.** This is a putative class action charging Google with copyright
21 infringement arising from the training of a number of generative AI text and image models
22 developed by disparate groups across the company. As this Court and others have recognized, “AI
23 cases in particular ... present a multitude of issues that require hard work and substantial
24 compromise to move forward with the litigation.” ECF No. 117 at 1; *In re OpenAI, Inc., Copyright*
25 *Infringement Litig.*, 2025 WL 1037221, at *1 (J.P.M.L. Apr. 3, 2025) (AI cases can involve
26 “complex[] and voluminous discovery regarding how defendants trained and designed their [AI
27 models]”). In the best of circumstances, discovery will be complicated and costly. But the burdens
28 will be made much worse if Plaintiffs do not engage in a thoughtful and cooperative process.

1 At Google's urging, and potentially to obviate some of the most burdensome (and
2 disproportionate) discovery, Judge Lee ordered early briefing on class certification and directed
3 the parties to prioritize discovery relevant to that issue. *See* ECF No. 93 at 21:19-21. Plaintiffs
4 thereafter amended their complaint to substitute in an improper fail-safe class definition. Google
5 moved to strike Plaintiffs' class allegations or stay non-class-related discovery. ECF No. 98. That
6 motion, and a separate motion to dismiss, are set for hearing before Judge Lee next week.

7 Despite Judge Lee's direction and Google's pending motions, Plaintiffs have pursued
8 extremely burdensome merits discovery (e.g., 260 Requests for Admission and 128 Requests for
9 Production, including requests for all "Training Data" for all of Google's generative AI models
10 and "[a]ll Documents and Communications related to the 'millions of content licenses that
11 authorize Google to use content [to] develop its products and services'") that would be
12 disproportionate in any case, but is especially so now. Google nevertheless has met and conferred
13 with Plaintiffs numerous times to ensure they have what they claim to need for class certification
14 and more.

15 Due to the enormous size of training data sets and the risks and burdens associated with
16 producing them, Google proposed that Plaintiffs select exemplary training data for review, with
17 appropriate stipulations to address concerns flowing from this reasonably-narrowed initial
18 approach. Google committed to facilitate Plaintiffs' informed selection of exemplary training data
19 through an efficient, collaborative process that is well underway. Far from "refus[ing] to provide
20 any meaningful information," ECF No. 120-2 at 3-4, Google agreed on March 26 that it would
21 make a substantial production of documents *identified by Plaintiffs* as "high-priority" beginning
22 April 11, with rolling productions to follow. Google has honored that agreement, producing over
23 1,000 documents on April 11 including extensive technical information relating to the generative
24 AI models at issue in this case, and will be producing additional documents describing training
25 data in depth this week. Google has also repeatedly offered to engage with Plaintiffs about training
26 data and answer Plaintiffs' reasonable questions through counsel following their review of the
27 materials. With the uncertain status of their class case coming before Judge Lee, Plaintiffs instead
28 rushed this motion forward, omitting the important context.

1 **Plaintiffs’ Deposition Notice.** Plaintiffs’ motion is addressed to a corporate deposition
2 notice they issued just after their first set of written discovery demands. When Google raised
3 objections, Plaintiffs proposed a 90-minute deposition on a single topic—“Training Data”—
4 encompassing the format, size, contents, and custodial issues of Google’s data, ostensibly to help
5 them select exemplary datasets for class certification. Hoping this might be a productive path
6 forward, Google asked Plaintiffs to explain what testimony they were looking for. Plaintiffs
7 responded by demanding testimony on 10 “subtopics” for the same 90-minute window. The
8 subtopics encompass nearly the entire case, including legal, policy, and technical issues such as
9 “Google’s doctrinal positions on fair use,” its licenses, the source and curation of datasets, and the
10 “technical architecture” of its AI systems. While Plaintiffs cannot meaningfully explore all 10
11 topics in 90 minutes, Google still must spend enormous time preparing a witness on their full scope
12 because Plaintiffs could choose to ask detailed questions on any of them. And the deposition, under
13 Plaintiffs’ proposal, would not count against their deposition limit.

14 In a March 20 meet and confer and in a March 21 letter, Google again objected, proposed
15 continuing its expedited production of high-priority documents and offered to answer targeted
16 questions to assist Plaintiffs in selecting representative training data for class certification.

17 **Argument.** Plaintiffs’ deposition demand is patently unreasonable. A 90-minute
18 deposition to cover 10 expansive topics is not practical, “targeted” or “proportional.” ECF No.
19 120-2 at 6; it is fanciful. Each topic encompasses enough material to require days of witness
20 preparation. But Plaintiffs could at best spend nine minutes on each—far less time than needed to
21 meaningfully cover the topics or justify the burdens of witness preparation. Moreover, several
22 topics have no bearing on Plaintiffs’ claimed purpose—understanding “the scope and content of
23 training data,” *id.*—and instead concern areas irrelevant to class certification or even proper Rule
24 30(b)(6) inquiry. That is, Plaintiffs have made no effort to propose targeted topics to assist their
25 understanding of training data.

26 The request is also premature. While it is reasonable for Plaintiffs to seek an
27 “understanding” of Google’s training data, ECF No. 120-2 at 5, this deposition is not the vehicle
28 to obtain it. Plaintiffs’ kitchen sink “subtopics” suggest they do not even know what they are

1 looking for. And the answers they would receive from witnesses forced to testify on such wide-
2 ranging topics with little time or direction to prepare are likely to be high-level and incomplete—
3 inviting further disputes over the adequacy of the witnesses. Indeed, that may be the point of the
4 demand.

5 Perhaps most problematic is Plaintiffs’ failure to adequately describe their subtopics.
6 Although Plaintiffs’ motion attaches a chart briefly listing them, they do not individually address
7 or explain the subtopics and did not include for the Court the somewhat more detailed descriptions
8 they provided Google, which highlight their impropriety. Plaintiffs’ “Fair Use” subtopic, for
9 example, seeks testimony on “Google’s doctrinal positions on fair use as applied to AI training
10 methodologies, comparing historical arguments with current positions regarding copyrighted
11 materials in AI datasets, and examining policies for identifying and excluding specific categories
12 of works based on fair use considerations.” Ex. 1 at 2. That subtopic, like many others, has nothing
13 to do with facilitating discovery by helping Plaintiffs understand and request Google’s training
14 data. Indeed, that topic would be improper in any context given that Plaintiffs are seeking Google’s
15 legal contentions. *3M Co. v. Kanbar*, 2007 WL 1794936, at *2 (N.D. Cal. June 19, 2007) (inquiries
16 seeking “legal conclusions ... should not form the basis for 30(b)(6) deposition topics”);
17 *Tradeshift, Inc. v. BuyerQuest, Inc.*, 2021 WL 2222811, at *2 (N.D. Cal. June 2, 2021) (topic
18 “improperly call[ed] for legal contentions from a lay witness”). Further, Rule 30(b)(6) requires
19 topics to be described with “reasonable particularity.” Fed. R. Civ. P. 30(b)(6). Plaintiffs’ proposed
20 subtopics, like the “[t]echnical architecture of Google’s AI systems” or Google’s “Data Usage
21 Policies And Licensing,” fail that test too, being so vague and expansive that no witness could
22 reasonably prepare to address them. *E.g.*, Ex. 1 at 2.

23 Rather than engage in a costly and unproductive deposition now, Google has offered a
24 practical alternative: produce high-priority documents on an expedited basis and continue working
25 cooperatively to provide information about training data informally. This approach will allow
26 Plaintiffs to review the high-priority documents, consult with their experts, and formulate focused
27 follow-up requests. If, after that process, a truly targeted Rule 30(b)(6) deposition is needed for
28 Plaintiffs to understand and select from Google’s training data, the Parties can likely agree on one.

1 But proceeding now with an unfocused, overbroad “preliminary” deposition while document
2 production is ongoing, will waste time, impose undue burdens, and generate unnecessary disputes.

3 For these reasons, Plaintiffs’ motion should be denied. If, for any reason, the Court permits
4 a deposition of Google now, it should serve as their lone corporate deposition for purposes of class
5 certification, and it should count against the deposition total they are allowed under the rules and
6 the scheduling order; they are not entitled to serial and unlimited depositions of Google. *See*
7 *Ramirez v. Bank of Am., N.A.*, 2024 WL 3012490, at *1 (N.D. Cal. June 13, 2024). Plaintiffs also
8 should not be permitted to seek to reopen the deposition on the basis of subsequently produced
9 documents when they made “a tactical decision” to demand an “early” deposition before document
10 production is close to complete. *See Bookhamer v. Sunbeam Prods. Inc.*, 2012 WL 5188302, at *3
11 (N.D. Cal. Oct. 19, 2012).

12 Respectfully submitted,

13 Dated: April 14, 2025

14 By: /s/ Paul J. Sampson

15 Paul J. Sampson, *pro hac vice*

16 Email: psampson@wsgr.com

17 **WILSON SONSINI GOODRICH & ROSATI**
Professional Corporation

18 95 S State Street, Suite 1000

19 Salt Lake City, UT 84111

20 Telephone: (801) 401-8510

21 David H. Kramer, SBN 168452

22 Email: dkramer@wsgr.com

23 Maura L. Rees, SBN 191698

24 Email: mrees@wsgr.com

25 **WILSON SONSINI GOODRICH & ROSATI**
Professional Corporation

26 650 Page Mill Road

27 Palo Alto, CA 94304-1050

28 Telephone: (650) 493-9300

Eric P. Tuttle, SBN 248440

Email: eric.tuttle@wsgr.com

WILSON SONSINI GOODRICH & ROSATI
Professional Corporation

701 Fifth Avenue, Suite 5100

Seattle, WA 98104-7036

Telephone: (206) 883-2500

Counsel for Defendants

GOOGLE LLC AND ALPHABET INC.